DOCUMENT RESUME

ED 156 214 IR 006 447

AUTHOR Evans, Richard M.; Surkan, Alvin J.
TITLE Effect of Computer-Delivered Testing on Achievement in a Mastery Learning Course of Study with Partial Scoring and Variable Pacing.
PUB DATE Jun 77
NOTE 8p.; Paper presented at the Conference on Computers in the Undergraduate Curricula (East Lansing, Michigan, June 20-22, 1977); For entire proceedings, see IR 006 142

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Achievement Gains; *Computer Oriented Programs; Educational Innovation; Educational Psychology; Educational Research; Evaluation Methods; *Mastery Learning; Pacing; Student Opinion; *Student Testing; Teacher Education Curriculum; Undergraduate Study
IDENTIFIERS *Computer Developed Instruction; Morningside College IA

ABSTRACT

The recent arrival of portable computer systems with high-level language interpreters now makes it practical to rapidly develop complex testing and scoring programs. These programs permit undergraduates access, at arbitrary times, to testing as an integral part of a mastery learning strategy. Effects of introducing the computer were studied by using, for control, data from students enrolled in the same undergraduate educational psychology course during the previous year. Results show that computer organized testing allowed students to take more replicates of each examination. Consequently there was, for three out of five examinations, an increase in the values of a measure of achievement based on the highest grade attained and time required before taking the final replicate desired for each examination of the course. Results of a student opinion survey indicated that in the presence of the computer organized testing, there was a greater feeling of learning than in a paper-and-pencil testing situation which allowed similar feedback with an invisible ink answer sheet. It was concluded that the feasibility of practical program development and that individual student utilization of portable, stand-alone, computer systems which now permit variable-paced mastery learning are propaedeutic in curriculum reform. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED156214

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

EFFECT OF COMPUTER-DELIVERED TESTING ON ACHIEVEMENT IN A MASTERY LEARNING COURSE OF STUDY WITH PARTIAL SCORING AND VARIABLE PACING

BY

Richard M. Evans

and

Alvin J. Surkan

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Ted Sjoersma

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

IR 006 447

Effect of Computer-Delivered Testing on Achievement in a Mastery Learning Course of Study With Partial Scoring and Variable Pacing

Richard M. Evans
Morningside College
Sioux City, Iowa
(712) 277-5821

Alvin J. Surkan
University of Nebraska
Lincoln, Nebraska

Abstract

The recent arrival of portable computer systems with high-level language interpreters now makes it practical to rapidly develop complex testing and scoring programs. These programs permit undergraduates access, at arbitrary times, to testing as an integral part of a mastery learning strategy. Effects of introducing the computer were studied by using, for control, data from students enrolled during the previous year. Results show that computer organized testing allowed students to take more replicates of each examination. Consequently there was, for three out of five examinations, an increase in the values of a measure of achievement based on the highest grade attained and time required before taking the final replicate desired for each examination of the course. Results of a student opinion survey indicated that in the presence of the computer organized testing, there was a greater feeling of learning than in a paper-and-pencil testing situation which allowed similar feedback with an invisible ink answer sheet. It was concluded that the feasibility of practical program development and that individual student utilization of portable, stand-alone, computer systems which now permit variable-paced mastery learning are propaedeutic in curriculum reform.

Introduction

The purpose of this study was to introduce a time-variable competency-based (TVCB) course of study into a liberal arts teacher education program and assess its effects on student achievement and attitudes. Earlier models of school learning suggested that maximum student achievement was provided for when the time allowed for learning could vary (Carroll, 1963), and teacher activities encouraged maximum pupil activities and pursuits (Harnischfeger and Wiley, 1975). It was hypothesized that the TVCB computer testing procedure stimulates more course-related student activity, with a consequent increase in student achievement, than does a similar traditionally taught course using a paper and pencil testing procedure.

Method

Cross-group contamination is a problem when introducing an experimental treatment into the curriculum of small schools, since most people know what others are doing. In the present case, achievement and attitude data available from a control group consisting of the previous year's educational psychology course, taught by the same instructor. Use of this control group cor-

responds to a version of something Tuckman (1972) calls a "Patched-Up" research design. These students had taken all five of Biehler (1974) examinations which accompany his text. They also had completed a test anxiety questionnaire (Evans, 1975), and a sixteen-question course evaluation questionnaire. The latter was designed by the instructor to record attitudes towards the course and mastery learning on a nine-point Likert-type bipolar scale.

Table 1 presents salient statistics available ascertaining the equivalence of the control group, which took educational psychology during the spring semester of 1975 and the experimental group, which took the course using the same Biehler materials one year later. Only the A.C.T. composite scores were gathered outside of the two classes Both groups were taught the course, subject to mastery learning ground-rules. These allowed students to retake, via replicates, any of the five course examinations in which they scored lower than they may have wished. A fixed criterion arbitrarily assigned A-B-C-D-F grades to intervals with lower bounds of 90-80-70-60-<60 percent, respectively. The control group was required to achieve a score of 90 percent or more on at least four of the five examinations to obtain an "A" for the course. However, a year later the intrinsic reinforcement of computer-delivered testing allowed the instructor to require that level of achievement on five out of five examinations.

Table 1

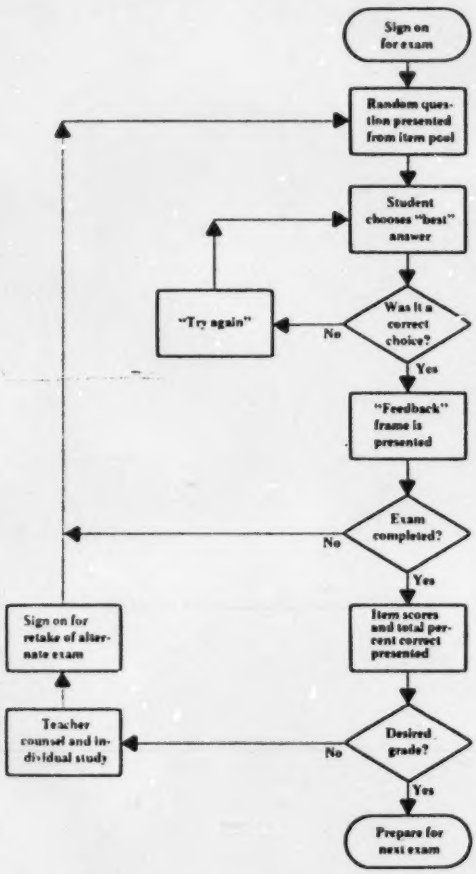
Measures Used to Equate Experimental and Control Groups

Group	A.C.T. Composite Score	Test Anxiety Score	Final Course Grade	Course Work Load	Time Required to Learn	Overall Course Rating
Experimental Group (Spring '76)						
X	21.86	11.76	3.24	5.30	4.49	6.78
s	5.05	5.76	.80	2.08	2.16	2.04
Control Group (Spring '75)						
X	22.76	12.22	1.96	5.47	4.82	6.20
s	5.12	6.41	1.11	1.51	2.01	2.18
Univariate F (df = 1,88)	0.67	0.12	1.74	0.20	0.57	1.64
p<	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.

Procedure

The treatment incorporated an IBM 5100 portable computer programmed to present replicate sets of fifteen questions randomly selected from each of the five Biehler test item pools. Students would make an appointment for a half hour out-of-class testing session. The computer was located in a carrell in a corner of the instructor's office. This made "counsel" following an examination a major aspect of the treatment. After signing on, students would be presented the randomly selected question. The question would be repeated until the examinee had identified the correct answer in an "answer until correct" procedure. After a "right" was flashed by the computer on the display screen, a feedback frame was then presented. This frame explained why one of the four multiple-choice responses was better than the other three, and referenced page numbers in the Biehler text. As the flow diagram of Figure 1 shows, students need not exit from the system until they achieve whatever mastery level their motivation allowed. The control group took their first attempt at each test in class, and

Figure 1. Morningside Mastery Strategy for Time-Variable Competency-Based Courses of Study.



their second and third attempts out of class. The examination papers utilized an invisible ink answer sheet. This also enabled students to keep selecting answers until an "X" signified that a correct choice had been made. The feedback frames were given at the end of the examinations. Both groups were given partial credit scores on a 3-2-1-0 basis, depending on the number of choices required to obtain the correct answer was 1,2,3, or 4 respectively. This scoring has the effect of increasing the effecting duration of the examination, slightly relax the fixed percentage all-or-nothing grade criterion, and increase the test reliability (Evans and Misfeldt, 1974). Since the control group had longer tests, consisting of twenty to thirty-five questions, the obtained partial scores were adjusted to fifteen question examinations to allow contrasts with the experimental group. Only the highest score on each of the five tests was recorded for grade purposes.

Subjects

The students were ninety sophomores enrolled during spring semester educational psychology courses in a private liberal arts college. Fifty-one students were in the 1975 (control) group and thirty-eight were in the 1976 (experimental) group. The numbers of these enrolled in music, elementary education, and secondary education were 11, 13, 27, and 5, 10, 23 in the two groups (Chi square = 1.09, df = 2, not significant). The groups did not differ significantly in A.C.T. composite scores, test anxiety scores, or final course grades in educational psychology. Since the testing method used in the course enabled most students to establish their expected course grades prior to the last week of the semester, three questions from the attitude measure were deemed important for equating the groups. They were: (1) Student perception of the course work load, (2) the time required for the students to learn compared to their perceptions of the amount required by other students, and (3) overall rating of the educational psychology course. As Table 1 shows, the groups did not differ on any of these measures.

Results

Measures opposite to the hypothesis were (a) scores on the five course examinations, (b) the number of attempts students made in order to achieve their particular course grade objective, and (c) student perception of the amount learned during the course. A total of 371 sets of examination questions were answered by the control group and 385 by the experimental group. Each student took an average of 7.27 and 9.87 examinations, respectively for each course. All measures were asceded univariate analysis of variance tests with a .05 level criterion for significance. As Table 2 reveals, the experimental group

Table 2

Measures Used to Determine Differences Between Experimental and Control Groups

Group	*No. Tries First Exam*	*No. Tries Second Exam*	*No. Tries Third Exam*	*No. Tries Fourth Exam*	*No. Tries Fifth Exam*	*Perception of Amount Earned*	*Scores Test One*	*Scores Test Two*	*Scores Test Three*	*Scores Test Four*	*Scores Test Five*
Experimental Group M=38											
$\bar{X}$	1.87	1.90	2.36	2.15	1.59	7.16	40.82	40.85	41.08	39.13	39.13
s	.77	.85	.84	1.23	.68	1.66	2.67	2.72	8.76	3.73	3.47
Control Group M = 51											
$\bar{X}$	1.41	1.35	1.80	1.57	1.14	5.80	40.45	38.55	37.65	38.10	37.51
s	.61	.59	.78	.78	.35	2.31	2.86	4.35	4.08	10.43	4.30
Univariate F											
(df=1,88)	10.11	12.75	10.51	7.60	16.96	9.32	.39	9.84	6.10	0.35	4.56
p	.01	.01	.01	.01	.01	.01	n.s.	.01	.02	n.s.	.05

took significantly more of each of all five examinations, and scored higher on three of them. This group also indicated learning more during the course than in regular courses relative to the control group.

The variable-pacing aspects of the experimental group were examined according to the number of 113 available days which elapsed between the semester beginning that each student achieved his target criterion level on each examination. Table 3 presents mean and standard deviations of these elapsed times for each examination. Most notable is the increased time variance near the middle or third examination, which was the most difficult experienced by the students.

Table 3

Mean and Standard Deviation of Days Taken Into Semester (Experimental Only)

	First Exam	*Second Exam*	*Third Exam*	*Fourth Exam*	*Fifth Exam*
$\bar{X}$	25.59	43.85	73.31	89.74	101.41
s	9.83	15.79	22.02	15.68	11.40

Discussion

The higher achievement in three of the course examinations by students conducted with the computer-organized procedure support Bloom's suggestion that most students can achieve at a level heretofore attainable only by the best students (Block, 1971). This higher achievement appears to be related to increased student activity outside of class, as shown by the increase in examinations taken. This being so, the key element of the Harnischfeger-Wiley model (1975), which asserts that activities which most affect "student pursuits" have the greatest direct effect on student achievement, is also supported by

the present findings. An earlier model for school learning (Carroll, 1962) suggested that if mastery were to be held constant, students would require variable periods of time to achieve that mastery. The consistent increases in standard deviations of elapsed time for achievement of the students' particular grade criterion until the third test, followed by a decrease in standard deviations of tests four and five, suggest however, that the present treatment did still not produce a completely time variable course of study. The constraints of the college's traditional fixed-time, variable-mastery curriculum required that after 113 days the semester end and letter grades be submitted to the registrar. Two students who had not taken all five examinations by that time were given no credit for the course.

The "threat" of no credit and the requirement to sign up for the course in another semester in order to complete its activities seemed to produce a flurry of activity during the last few weeks of the semester, possibly causing the decrease in standard deviation of lapsed time in taking the last two examinations. In this sense, one of the major obstacles to a "pure" mastery learning strategy—the fixed time units in the semester—was assaulted on an individual course basis, but not yet conquered because the remainder of the college curriculum shaped student attitudes toward finishing when other courses ended, rather than opting to repeat segments of the course not completed at the 90 percent level.

The increase in examination scores without a corresponding increase in course grades should be mentioned. Again, the instructor felt that the experimental treatment allowed him to increase the grade requirement to five out of five for a particular grade. This increase in stringency was not felt possible using the paper-and--pencil testing procedure.

The present findings are in agreement with a study by Cartwright and Derevensky (1975) who also used computer-organized testing with the Biehler educational psychology materials. Their attitudinal study found that students perceived this to be an effective method to incorporate in mastery of course material.

It should be noted that the feedback following each of the multiple-choice questions and the counsel by the instructor or his assistant following examinations were an important part of the experimental treatment. Such human contacts tended to personalize what could become a rather mechanistic, impersonal, and private method of individually progressing through the material.

Indeed, Block and Burns (1976) have attempted to show the emphasis of mastery learning on the social aspects of education over the relatively private individually prescribed instruction. Office hours during the experimental period were increased by a factor of four in order to discuss test questions and other educational psychology subject matter. A significant portion of this time was spent trying to convince some students that they did have the ability to score well on these tests, and to keep trying. The immediate feedback often revealed that students were missing material which they said they knew quite well. This seemed to be due to poor reading habits, which improved as they acquired *savoir faire* in taking tests. Other students remarked that they enjoyed the text book more than had others. It was conjectured that these students, for perhaps the first time in their academic careers, were required to read technical material to mastery, which in turn brought pleasure. Finally, the social aspects of the treatment in both groups was enjoined by the requirement that all students attend all class discussions, no matter how far they had progressed individually through the material.

Conclusion

With the arrival of portable computers with high-level language interpreters, small schools can now enter the computer-assisted instruction arena, which has in the past been limited to larger universities possessing the

large computers. The time-variable-competency-based course using the computer-delivered testing allowed by low cost portable computers shows promise in providing methods for keeping track of a multitude of individual student differences in time required for learning. This is seen as the greatest obstacle to implementation of pure forms of mastery learning in the schools. If the basic theory of mastery learning is sound, then the time has come that we reorganize schools to be in accord with this approach rather than attempt to modify our mastery learning strategies to fit into the existing school systems. The results of the present experiment have demonstrated that this time is at hand since the technology is already available.

References

Biehler, R. *Instructor's Manual to Accompany Psychology Applied to Teaching.* Boston: Houghton Mifflin Co., 1974.

Block, J. H. *Mastery Learning. Theory and Practice.* N.Y.: Holt, Rinehart and Winston, Inc., 1971.

Block, H. H. and Burns, R. B. Mastery Learning. Mimeographed. University of California, Santa Barbara, 1976.

Carroll, J. B. "A Model for School Learning." *Teachers College Record.* 1963, 64, 723-733.

Cartwright, G. F., and Derevensky, J. L. An Attitudinal Study of Computer Assisted Testing as a Learning Method. Paper presented to the 59th Annual Meeting of the American Educational Research Association. Washington, D.C., April 1975.

Evans, R. M. Confidential Questionnaire Concerning Education 215. Mimeographed. Morningside College, Sioux City, Iowa, 1975.

Evans, R. M., and Misfeldt, K. Effect of self-scoring procedures on test reliability. *Perceptual and Motor Skills.* 1974, 38, 1248.

Harnischfeger, A., and Wiley, D. E. "Teaching-Learning Processes in Elementary School: A Synoptic View." *Studies of Educative Processes.* No. 9. University of Chicago, 1975.

Tuckman, B. W. *Conducting Educational Research.* New York: Harcourt, Brace, Jovanovich, Inc. 1972.